

**TELEMATIKA**, Vol. 15, No. 01, APRIL, 2018, Pp. 77 – 86  
**ISSN 1829-667X**

## **IDENTIFIKASI KEBUTUHAN DASAR DI TEMPAT EVAKUASI SEMENTARA PASCA ERUPSI MERAPI DENGAN SENTIMENT ANALISIS DAN SUPPORT VECTOR MACHINE**

**Resky Rayvano Moningka<sup>(1)</sup>, Djoko Budiyo Setyohadi<sup>(2)</sup>, Khaerunnisa<sup>(3)</sup>, Pranowo<sup>(4)</sup>**

Jurusan Teknik Informatika, Universitas Atma Jaya Yogyakarta

Gedung Bonaventura, Caturtunggal, Depok, DIY, 55281, INDONESIA<sup>(1,2,4)</sup>

Jurusan Teknik Arsitektur, Universitas Atma Jaya Yogyakarta

Gedung Thomas Aquinas, Caturtunggal, Depok, DIY, 55281, INDONESIA<sup>(3)</sup>

reskym@gmail.com<sup>(1)</sup>, djoko.bdy@gmail.com<sup>(2)</sup>, khaerunnisa@staff.uajy.ac.id<sup>(3)</sup>,

pran@mail.uajy.ac.id<sup>(4)</sup>

### **Abstract**

*Mount Merapi Eruption in 2010 was the biggest after 1872. The impact of this eruption was felt by people who lived around the areas which were affected by this Merapi Eruption. Thus, disaster management was done. One of the disaster management was the fulfillment of basic needs. This research aims to collect public opinion against the fulfillment of basic needs in the shelters after Merapi Eruption based on Twitter data. The algorithm which is used in this research is Support Vector Machine to develop classification model over the data that has been collected. The expected result from this study is to know the basic needs in a shelter. The accuracy gained by performing Cross Validation for 10 folds from Support Vector Machine is 87.96% and Maximum Entropy is 87.45%.*

**Keywords:** *twitter, sentiment analysis, merapi eruption, support vector machine*

### **Abstrak**

Erupsi Gunung Merapi 2010 merupakan yang terbesar setelah tahun 1872. Dampak dari Erupsi Gunung Merapi dirasakan oleh masyarakat yang tinggal di daerah terdampak Erupsi Merapi. Oleh sebab itu dilakukan penanggulangan Bencana. salah satu penanggulangan bencana adalah pemenuhan kebutuhan dasar. Penelitian ini bertujuan untuk mengumpulkan opini publik terhadap pemenuhan kebutuhan dasar di tempat pengungsian pasca erupsi merapi berdasarkan data Twitter. Algoritma yang digunakan dalam penelitian ini adalah Support Vector Machine untuk membangun model klasifikasi atas data yang sudah dikumpulkan. Hasil yang diharapkan dari penelitian ini adalah mengetahui kebutuhan dasar dari suatu tempat pengungsian. Akurasi yang didapatkan dengan melakukan Cross Validation sebanyak 10 fold dari model klasifikasi Support Vector Machine 87,96% dan Maximum Entropy 87,45

**Kata Kunci:** *twitter. analisis sentimen. erupsi merapi. support vector machine*

## **1. PENDAHULUAN**

Bencana merupakan peristiwa yang mana menjadi ancaman bagi kehidupan manusia. Bencana dapat dikategorikan menjadi dua, faktor alam dan non alam. Menurut Kepmen No.17/kep/Menko/Kesra/x/95 bencana adalah rangkaian peristiwa yang disebabkan oleh alam, manusia, dan atau keduanya yang mengakibatkan korban dan penderitaan manusia, kerugian harta benda, kerusakan lingkungan, kerusakan sarana prasarana dan fasilitas umum serta menimbulkan gangguan terhadap tata kehidupan dan penghidupan masyarakat

Menurut UU 24/2007, Penanggulangan bencana adalah serangkaian upaya yang meliputi penetapan kebijakan pembangunan yang beresiko timbulnya bencana, kegiatan pencegahan bencana, tanggap darurat, rehabilitasi dan rekonstruksi. Berdasarkan pasal 31, ada 4 aspek penyelenggaraan penanggulangan bencana salah satunya sosial, ekonomi, dan budaya

masyarakat. Pada pasal 33, ada 3 tahap penyelenggaraan penanggulangan bencana yaitu prabencana, saat tanggap darurat, dan pascabencana. Pada tahap tanggap darurat, berdasarkan pasal 48 d pemenuhan kebutuhan dasar meliputi kebutuhan air bersih, pangan, sandang, pelayanan kesehatan, pelayanan kesehatan, pelayanan psikososial dan penampungan dan tempat hunian. Penulis melakukan penelitian sentiment analysis untuk mengetahui kebutuhan dasar yang ada di tempat pengungsian.

*Sentiment Analysis* merupakan metode analisis yang dipakai untuk mengidentifikasi tentang komentar para pengguna. Menurut Chintala (2012) *sentiment analysis* merupakan metode menganalisis sepotong data untuk emosi manusia. Menurut (GO, Huang, Bhayani, 2009), *sentiment analysis* merupakan area penelitian yang menonjol dan aktif yang didorong dengan pesatnya pertumbuhan media sosial dan kesempatan untuk mengakses opini berharga dari banyak kalangan masalah bisnis, dunia, dan sosial. Pada kesempatan ini, penulis melakukan *sentiment analysis* berdasarkan data *tweet* pengguna media sosial twitter.

Twitter adalah layanan *microblogging* yang diluncurkan secara formal pada tanggal 13 Juli 2006 (Mustafa 2013). Twitter dianggap *microblog* karena pusat aktivitas berkisar seputar posting update singkat atau *tweets*. Ukuran maksimal blog adalah 140 karakter. Twitter merupakan *microblog* yang menjadi sumber yang hampir tak terbatas dalam teks klasifikasi (Torunoğlu, Telseren, Sağtürk, dan Ganiz, 2013). Dalam melakukan *sentiment analysis*, ada beberapa algoritma-algoritma *Machine Learning* yang dapat dipakai, dalam hal ini penulis menggunakan algoritma *Support Vector Machine*.

*Support Vector Machine* (SVM) diperkenalkan Vapnik, Boser, Guyon 1992, kumpulan kombinasi dari teori-teori komputasi seperti margin hyperlane (duda & hurt tahun 1973), kernel (Aronszajn 1950). SVM merupakan metode pembelajaran linier yang menemukan *hyperlane optimal* untuk memisahkan dua kelas (positif dan negatif). Menurut Mornes et al (2013, 624) telah banyak penelitian melaporkan bahwa SVM merupakan metode yang paling akurat. Keistimewaan dari SVM adalah kemampuan dalam menerapkan pemisah input data non-linier berdimensi tinggi. Itu diperoleh dengan menggunakan fungsi kernel yang diperoleh.

Elly Indrayuni (2016) melakukan analisa sentimen review hotel menggunakan algoritma *Support Vector Machine* berbasis Particle Swarm Optimization. Menurut peneliti, hotel merupakan salah satu produk pariwisata yang sangat penting untuk dipertimbangkan baik dari segi fasilitas, pelayanan, ataupun jarak tempuh perjalanan wisata. Dengan munculnya berbagai macam website wisata yang menyediakan fasilitas bagi pengguna internet untuk menuliskan opini dan pengalaman pribadi secara online. Dengan banyaknya opini yang diberikan oleh pengguna, maka tidak dimungkinkan untuk dibaca secara keseluruhan untuk mengambil suatu keputusan. Maka dari itu dilakukan analisis sentimen. Dalam penelitian ini, metode yang dipilih adalah *Support Vector Machine*. Alasan peneliti memilih metode ini karena SVM mampu mengidentifikasi *hyperplane* terpisah yang memaksimalkan margin antardua kelas yang berbeda. Serta keistimewaan SVM yang dapat menerapkan pemisah linier pada input data non linear berdimensi tinggi dengan menggunakan fungsi kernel. Peneliti menggunakan SVM dan Particle Swarm Optimization (PSO) sebagai seleksi fitur untuk meningkatkan nilai akurasi analisa sentimen. Penelitian ini menggunakan 300 data review hotel yang terdiri dari 150 data opini positif dan 150 data opini negatif berdasarkan review terbaru dari situs [www.tripadvisor.com](http://www.tripadvisor.com). Dengan menggunakan metode SVM menghasilkan akurasi sebesar 91.33%. Dan penerapan dengan menggunakan seleksi fitur PSO pada algoritma SVM membuat nilai akurasi meningkat menjadi 96.94%.

Tiara, Mira Kania Sabariah, Veronikha Effendy (2015) *Sentiment Analysis in Twitter Using Combination of Lexicon-Based and Support Vector Machine for Assessing the Performance of a Television Program* membahas tentang peningkatan kualitas dari sebuah program televisi. Komentar dari pengguna Twitter dapat melengkapi penilaian program televisi yang biasanya dilakukan dengan menggunakan rating dan terwakilkan dalam bentuk kuantitas. Dengan banyaknya komentar pengguna Twitter, diharapkan dapat melengkapi asesmen kualitas. Dalam melakukan analisis sentimen, mereka memakai metode SVM yang dikombinasi dengan

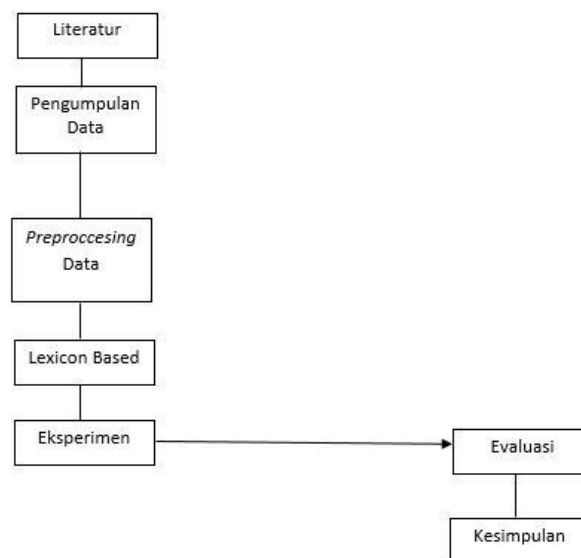
Lexicon-Based. Kedua metode ini memiliki karakteristik yang berbeda, tetapi dapat saling melengkapi. Metode Lexicon dipakai untuk membuat label tweets yang dijadikan data training di

SVM, sehingga tidak akan ada proses pelabelan secara manual. Penelitian ini menggunakan lima program televisi sebagai data penelitian. Data dipilih secara acak dan dibagi menjadi data *training* dan data *testing*. Total tweets dari lima program ini adalah 2200 tweet yang terbagi atas program A 300 tweets, program B 300 tweets, program C 500, program D 500, dan program E 600. Ada beberapa skenario yang dilakukan dalam pengujian perbandingan, perbandingan kedua data sebagai berikut:

1. Data pelatihan 60%: Data uji 40%
2. Data pelatihan 70%: Data uji 30%
3. Data pelatihan 80%: Data uji 20%
4. Data pelatihan 90%: Data uji 10%

dengan beberapa skenario diatas, nilai rata-rata tertinggi untuk akurasi ada pada perbandingan data pelatihan 90% dan data uji 10%. Setelah semua proses selesai dilakukan, didapatkan program televisi yang didominasi oleh sentimen positif merupakan salah satu yang memiliki nilai keakuratan yang tinggi sebesar 80%.

## 2. METODOLOGI PENELITIAN



Gambar 1. Metodologi Penelitian

### 2.1. Pengumpulan Data

Pada tahap ini, penulis melakukan pengumpulan data yang digunakan untuk eksperimen. Penulis memakai library *jefferson-Henrique/GetOldTweets* pada perangkat lunak Netbeans untuk *grabbing* data dari twitter. *GetOldTwitter* (GOT) Java merupakan program pencarian data pada Twitter yang ditulis oleh Jefferson-Henrique (Setyaji, B.W., 2017). Dalam pengumpulan data, penulis memakai kata kunci yang berkaitan dengan tema penelitian. Setelah data dikumpulkan, tahap selanjutnya adalah *preprocessing* data.

### 2.2. Preprocessing Data

*Preprocessing* data merupakan tahap pembentukan data yang ideal yang akan digunakan dengan optimal pada metode klasifikasi. *Preprocessing* data adalah proses pembersihan dan mempersiapkan teks untuk klasifikasi (Hadid et al, 2013). Ada beberapa tahap yang dilakukan dalam *preprocessing*, tahapan yaitu Case Folding, Penghilangan Karakter HTML, ikon emosi,

hashtag(#), url, tanda baca, symbol, dll, serta Stopwords untuk penghilangan kata tidak bermakna dan kata imbuhan.

### 2.3. Pemrosesan Data

#### 2.3.1. Lexicon Based

*Lexicon based* adalah teknik *unsupervised*, metode ini mengklasifikasikan data ke dalam dua kelas, kelas positif atau negatif. Metode ini berbasis *lexicon* ini menggunakan bantuan kamus untuk mengklasifikasikan *tweet* menjadi sentiment positif atau negatif. Ada beberapa langkah berdasarkan *lexicon* itu digunakan seperti menentukan polaritas kata-kata, penanganan negasi, dan juga memberikan skor untuk setiap masing-masing entitas dalam *tweet*. Rumus untuk menghitung nilai dari suatu entitas sebagai berikut

$$\text{Score}(e) = \sum_{w_i: w_i \in L \cap w_i \in s} \frac{w_i \cdot s_o}{\text{dis}(w_i, e)} \quad (1)$$

Score(e) = nilai akhir suatu entitas

W<sub>i</sub> = kata pendapat

L = semua kata pendapat

s = kalimat yang berisi kata-kata entitas dan Opini

s<sub>o</sub> = label kata pendapat (+1 or -1)

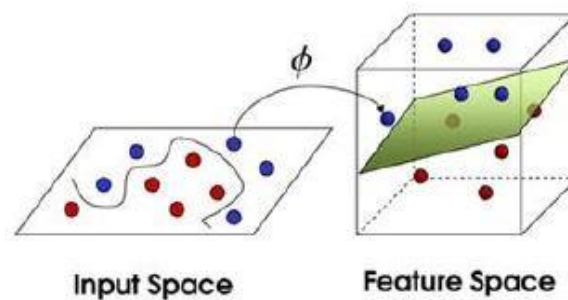
dis(w<sub>i</sub>, e) = jarak entitas(e) dan opini (w<sub>i</sub>)

#### 2.3.2. Pembentukan Model

Tahap ini bertujuan untuk mendapatkan parameter-parameter yang ideal dalam proses pengklasifikasian data. Dengan menggunakan algoritma SVM dengan kernel (non linear), maka diperlukan beberapa proses berikut:

##### Pemilihan Kernel

Terdapat beberapa kernel umum atau sering digunakan dalam metode SVM. Kernel tersebut adalah linier, *Radial Basis Function* (RBF), polynomial, dan sigmoid. jenis kernel tersebut yang paling umum digunakan adalah RBF. RBF memiliki sifat yaitu dapat menangani permasalahan *soft-margin classifier*. RBF memiliki parameter C sebagai penentu besarnya penalti akibat kesalahan klasifikasi.



Gambar 2. Penggunaan teknik kernel untuk merubah *input space* menjadi *feature space* (Haltuf, 2014)

##### Menentukan Nilai Variabel Cost dan Gamma

Dalam penelitiannya, (Haltuf, 2014) menggunakan rentang nilai C = 2<sup>-5</sup>, 2<sup>-3</sup>, ..., 2<sup>15</sup> dan gamma = 2<sup>-15</sup>, 2<sup>-13</sup>, ..., 2<sup>3</sup>. Adapun pada libSVM, diperlihatkan nilai yang sama. Penulis akan menggunakan nilai C dan gamma yang sama dalam nilai rentangan tersebut.

### 2.4. Interpretasi Data

Pada tahap ini, penulis melakukan pemetaan terhadap tempat pengungsian ydari daerah yang terdampak erupsi gunung Merapi seperti Kabupaten Sleman, Kabupaten Klaten, Kabupaten

Boyolali, dan Kabupaten Magelang. Setelah memetakan tempat pengungsi, penulis melakukan identifikasi terhadap kebutuhan serta permasalahan dari tempat pengungsi tersebut. Setelah berhasil mengidentifikasi kebutuhan serta permasalahan, penulis mengidentifikasi sentiment dari kebutuhan dan permasalahan tersebut .

## 2.5. Evaluasi Data

Setelah dilakukan preprocessing, pemrosesan data, pembentukan model dan terbentuk model klasifikasi akan dilakukan evaluasi. Evaluasi ini bertujuan untuk menentukan baik atau tidaknya model yang terbentuk. Evaluasi model apada penelitian ini menggunakan RStudio.

### 2.5.1. Analisis Model dengan R Language (RStudio)

Pada tahap ini, penulis akan membahas hasil dari RStudio. Pembahasan tersebut meliputi *confusion matrix* yang terbentuk, *mean absolute error* (MAE), nilai area under curve (AUC) pada kurva *Receiver Operating Characteristic*(ROC), k-fold cross validation. *Confusion matrix* merupakan matriks yang berisi tentang seberapa besar valid klasifikasi yang dilakukan oleh suatu metode klasifikasi. Data yang diberikan berupa nilai *True Positive*(TP), *False Positive*(FP), *True Negative*(TN), *False Negative*(FN).

TN merupakan jumlah data yang diprediksi masuk ke kelas negatif dan memiliki kelas aktual dikelas negatif. FP merupakan jumlah data yang diprediksi masuk ke kelas positif tetapi memiliki kelas aktual dikelas negatif. FN merupakan jumlah data yang diprediksi masuk ke kelas negatif tetapi memiliki kelas aktual dikelas positif. TP merupakan jumlah data yang diprediksi masuk ke kelas positif dan memiliki kelas aktual dikelas positif.

TP dan TN menunjukkan keberhasilan metode yang digunakan dalam memprediksi data. FP dan FN menunjukkan kesalahan metode yang digunakan dalam memprediksi data. Dari confusion matrix, dapat dicari nilai *accuracy*, *precision*, *recall*. Formulasi untuk *accuracy* sebagai berikut:

$$\text{Accuracy} = \frac{TN+TP}{TN+TP+FP+FN} \quad (2)$$

Formula untuk *Precision* sebagai berikut:

$$\text{Precision} = \frac{TP}{TP+FP} \quad (3)$$

Formula untuk *Recall* sebagai berikut:

$$\text{Recall} = \frac{TP}{TP+FN} \quad (4)$$

Dari nilai-nilai diatas, penulis memvisualisasikan akurasi dengan menggunakan ROC. *Receiver Operating Characteristic* (ROC) adalah teknik untuk menggambarkan, mengorganisasikan dan memilih pengklasifikasi berdasarkan kinerja mereka (Fawcett, 2006). Kurva ROC digunakan untuk mengukur nilai *Area Under Curve* (AUC).

Pedoman untuk mengklasifikasikan keakuratan pengujian menggunakan nilai AUC, sebagai berikut (Gorunescu, 2011):

- a. 0.90-1.00 = Excellent Classification
- b. 0.80-0.90 = Good Classification
- c. 0.70-0.80 = Fair Classification
- d. 0.60-0.70 = Poor Classification
- e. 0.50-0.60 = Failure

## 2.6. Perbandingan SVM dengan metode lain

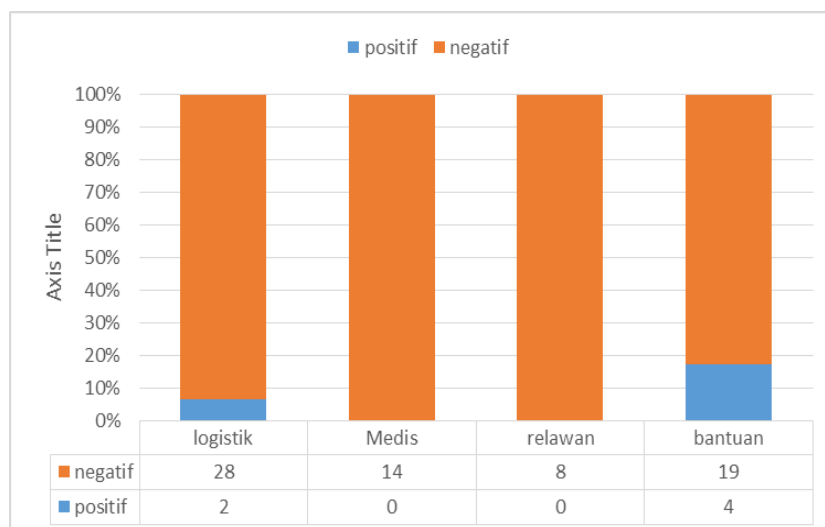
Proses ini bertujuan untuk melihat apakah metode klasifikasi SVM lebih baik dari metode klasifikasi yang lainnya dengan melihat precision, recall, fmeasure, dan akurasi yang didapatkan.

## 2.7. Kesimpulan

Pada tahap ini, akan dilakukan penarikan kesimpulan setelah melakukan berbagai tahap penelitian. Kesimpulan ini dilihat dari nilai akurasi yang didapatkan apakah baik dalam mengklasifikasikan data dibandingkan dengan metode lainnya. Hasil yang didapatkan juga akan dilakukan diskusi apakah manfaat dari penelitian ini dan membahas tentang kebutuhan dasar yang sudah teridentifikasi.

## 3. HASIL PEMBAHASAN

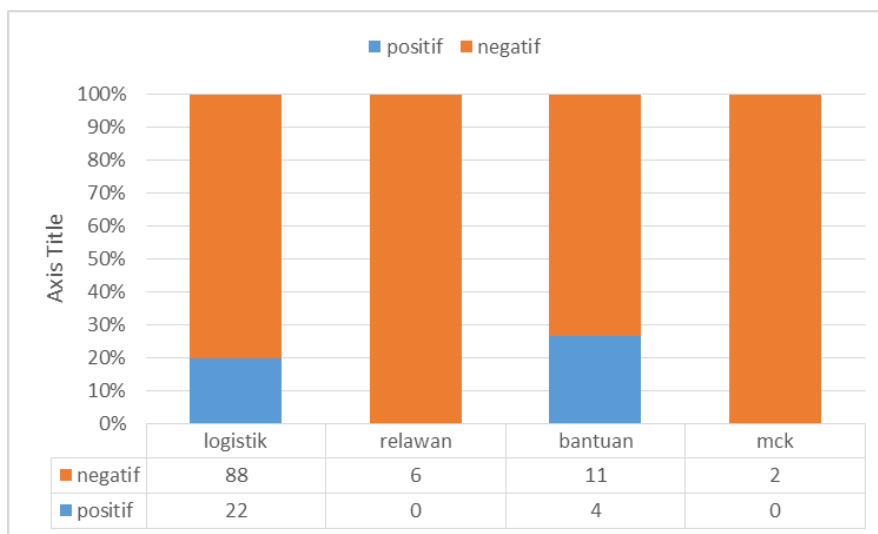
Setelah penulis melakukan pengumpulan data, melakukan preprocessing data, pembobotan, dan melakukan evaluasi, maka data siap untuk diinterpretasikan. Pada tahap ini, penulis memisahkan data berdasarkan daerah yang terkena dampak erupsi gunung Merapi yaitu Kabupaten Sleman, Kabupaten Klaten, Kabupaten Boyolali, dan Kabupaten Magelang. Setelah memisahkan data, penulis mengidentifikasi tempat pengungsian berdasarkan masing-masing daerah tersebut. Setelah itu, penulis melakukan pemetaan terhadap tempat pengungsian berdasarkan *tweet* masyarakat. Setelah melakukan pemetaan, penulis melakukan identifikasi terhadap kebutuhan dasar dari tempat pengungsian tersebut berdasarkan *tweet* yang dipost masyarakat melalui media sosial Twitter. Dari kebutuhan tersebut, penulis mengidentifikasi sentiment positif dan sentiment negatif dari kebutuhan dasar. Setelah mengetahui sentiment positif dan sentiment negatifnya, penulis mengidentifikasi kata *tweet* yang mewakili opini negatif dan kata *tweet* yang mewakili opini positif. Tempat pengungsian yang akan diidentifikasi tersebar di beberapa lokasi sebagai berikut:



**Gambar 3. Grafik Sentiment dari tweet Kebutuhandasar Kabupaten Sleman**

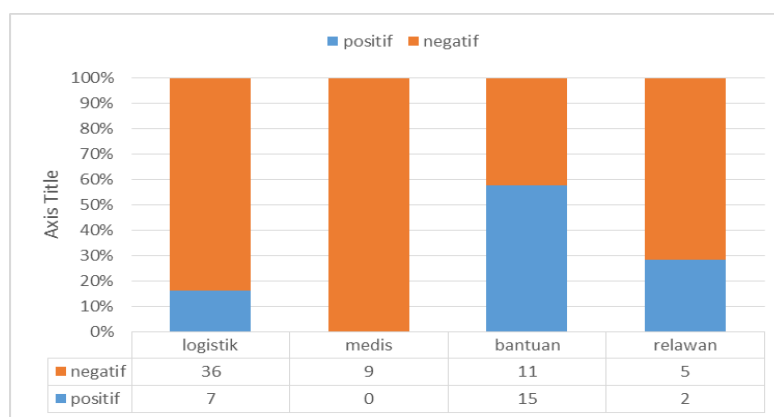
Dari Gambar 3 dapat dilihat kebutuhan dasar yang diperlukan untuk tempat pengungsian di Kabupaten Sleman. Kebutuhan dasar tersebut meliputi Kebutuhan Logistik, Kebutuhan Medis, Relawan, dan Bantuan. Data kebutuhan dasar diatas dikumpulkan berdasarkan tempat pengungsian yang berhasil diidentifikasi di wilayah Kabupaten Sleman. Tempat pengungsian yang teridentifikasi berdasarkan *tweet* adalah 15 titik tempat pengungsian yang tersebar di 8 kecamatan. Dari *tweet* yang didapatkan, opini negatif yang berhasil diidentifikasi meliputi butuh, meninggal, mengeluh, kurang, lambat, belum tersentuh. Sedangkan *tweet* yang mewakili opini

positif meliputi mencukupi, dalam perjalanan, kirim, menyiapkan. Dari grafik diatas, dapat dilihat bahwa sentiment di Kabupaten Sleman didominasi oleh *tweet* negatif.



**Gambar 4. Grafik Sentiment dari tweet kebutuhandasar Kabupaten Klaten**

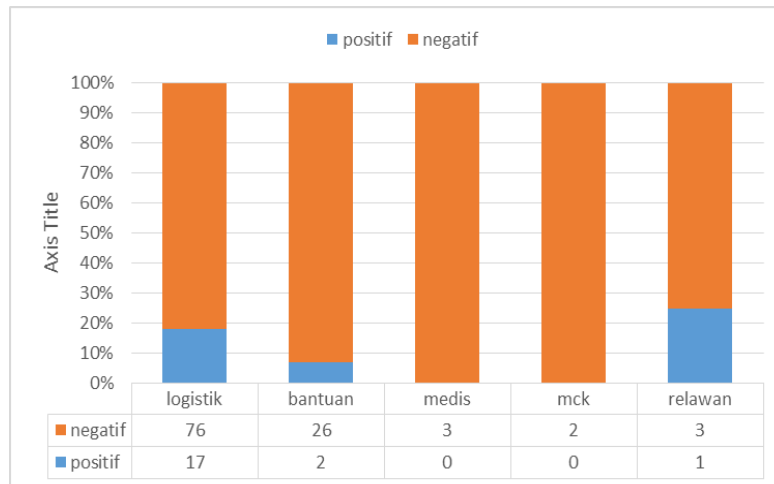
Dari Gambar 4 dapat dilihat kebutuhan dasar yang diperlukan untuk tempat pengungsian di Kabupaten Klaten. Kebutuhan dasar tersebut meliputi Kebutuhan Logistik, Relawan, Bantuan, MCK. Data Kebutuhan diatas dikumpulkan berdasarkan tempat pengungsian di wilayah Kabupaten Klaten yang teridentifikasi. Tempat pengungsian yang teridentifikasi berdasarkan *tweet* adalah 15 titik tempat pengungsian yang tersebar di 8 kecamatan. *Tweet* yang mewakili opini negatif meliputi penimpunan, butuh, kurang, mohon, belum tersentuh. Sedangkan *Tweet* yang mewakili opini positif meliputi kirim, terima, memilih, kelebihan, bawa. Dari grafik diatas, sentiment di Kabupaten Klaten didominasi dengan *tweet* negatif.



**Gambar 5. Grafik Sentiment dari tweet kebutuhan dasar Kabupaten Boyolali**

Dari Gambar 5 dapat dilihat kebutuhan dasar yang diperlukan untuk tempat pengungsian di Kabupaten Boyolali. Kebutuhan dasar tersebut meliputi Kebutuhan Logistik, Medis, Bantuan, Relawan. Data Kebutuhan diatas dikumpulkan berdasarkan tempat pengungsian di wilayah Kabupaten Boyolali yang teridentifikasi. Tempat pengungsian yang teridentifikasi berdasarkan *tweet* adalah 13 titik tempat pengungsian yang tersebar di 7 kecamatan. *Tweet* yang mewakili opini negatif meliputi belum tersentuh, meninggal, butuh, minim. Sedangkan *Tweet*

yang mewakili opini positif meliputi terima, bergerak, kirim, melimpah. Dari grafik diatas, sentiment di Kabupaten Boyolali didominasi dengan *tweet* negatif. Namun berdasarkan grafik diatas, dari sisi Bantuan Kabupaten Boyolali memiliki presentase yang didominasi *tweet* positif tetapi tidak teridentifikasi secara detail bantuan apa yang diperlukan.



**Gambar 6. Grafik Sentiment dari tweet kebutuhan dasar Kabupaten Magelang**

Dari Gambar 6. dapat dilihat kebutuhan dasar yang diperlukan untuk tempat pengungsian di Kabupaten Magelang. Kebutuhan dasar tersebut meliputi Kebutuhan Logistik, Bantuan, Medis, MCK, Relawan. Data Kebutuhan diatas dikumpulkan berdasarkan tempat pengungsian di wilayah Kabupaten Magelang yang teridentifikasi. Tempat pengungsian yang teridentifikasi berdasarkan *tweet* adalah 12 titik tempat pengungsian yang tersebar di 11 kecamatan. *Tweet* yang mewakili opini negatif meliputi butuh, kurang, nyuri, tolong. Sedangkan *Tweet* yang mewakili opini positif meliputi nyortir, kirim. Dari grafik diatas, sentiment di Kabupaten Boyolali didominasi dengan *tweet* negatif.

### Evaluasi Model

Pada bagian ini, penulis akan mengevaluasi pengklasifikasian dari metode SVM. Proses yang akan dijelaskan adalah analisa Kurva ROC, AUC, dan akurasi. Dengan menggunakan nilai  $C = 2^9$  dan  $\gamma = 2^{-1}$ , mendapatkan nilai precision, recall, fmeasure sebagai berikut:

**Tabel 1 Tabel Akurasi SVM dalam mengklasifikasi**

Tweet	14626
PRECISION	89%
RECALL	84,3%
FMEASURE	86,3%

Dari tabel 1 dapat diketahui bahwa akurasi yang didapatkan dengan menggunakan variabel  $C = 2^9$  dan  $\gamma = 2^{-1}$  mendapatkan akurasi sebesar 86,3%. Setelah mendapatkan akurasi, penulis melakukan *k-fold cross validation* untuk mendapatkan nilai akurasi yang lebih baik. Penulis melakukan *cross validation* sebanyak  $k=10$ . Berikut hasil dari  $k=10$

**Tabel 2. Tabel K-fold validation**

K-1	85,97%
K-2	85,94%
K-3	86,82%
K-4	84.52%



K-5	85,95%
K-6	85,71%
K-7	87,15%
K-8	87,96%
K-9	85,70%
K-10	84,93%

### Perbandingan SVM dengan metode lain

Dari proses yang sudah dilalui, pada tahap ini penulis melakukan perbandingan antara metode SVM dan metode yang lainnya. Pada penelitian ini, penulis membandingkan dengan metode *Maximum Entropy*. Penulis melakukan identifikasi terhadap precision, recall, dan fmeasure dari *Maximum Entropy* dan membandingkan hasil dengan SVM.

**Tabel 3. Perbandingan Metode**

Metode	Precision	Recall	Fmeasure	Akurasi
SVM	0,890	0,843	0,863	87,96%
MAXENT	0,823	0,803	0,810	87,45%

Dari Tabel 3. dapat dilihat bahwa hasil penelitian yang dilakukan, dengan melakukan k-fold validation sebanyak k=10, nilai akurasi metode SVM lebih tinggi dibandingkan metode MAXENT. Dapat dilihat pula dari nilai precision, recall, dan fmeasure SVM memiliki nilai lebih besar dibandingkan dengan MAXENT. Maka dapat disimpulkan bahwa pada penelitian ini metode SVM lebih baik dalam pengklasifikasian data dibanding dengan metode MAXENT.

### KESIMPULAN

Kebutuhan dasar dalam penanganan bencana merupakan hal yang penting untuk diketahui sebagai data bagi pengelolaan bencana. Dalam penelitian kami terlihat bahwa teridentifikasi perlunya kebutuhan dasar di tempat pengungsian berdasar sentiment negative yang meliputi : kebutuhan Logistik, kebutuhan Relawan, pelayanan kesehatan, kebutuhan MCK. Dengan penggunaan *k-fold cross validation*, dapat disimpulkan bahwa penggunaan sentiment analisis memakai algoritma SVM dan kernel Radial Basis Function sudah memenuhi akurasi yang didapatkan nilai akurasi yang sebesar 87,96% dengan nilai AUC rata-rata diantara 0,877-0,891

### DAFTAR PUSTAKA

- Anto, N.S., A. B., & Handoko. (2003). Support Vector Machine -Teori dan Aplikasinya dalam Bioinformatika-. dari <http://ilmukomputer.com>.
- Banados, J. A., & Espinosa, K. J. (2014). Optimizing support vector machine in classifying sentiments on product brands from Twitter. In *Information, Intelligence, Systems and Applications, IISA 2014, The 5th International Conference on* (pp. 75-80)
- BNPB. (2017, November 17). *Siaga Bencana*. Diambil kembali dari Badan Nasional Penanggulangan Bencana: <https://www.bnpb.go.id/home/definisi>
- BNPB 2010a. (2010). Peta rekapitulasi korban, pengungsi dan kerusakan akibat letusan gunung api Berapi. *Badan Nasional Penanggulangan Bencana*.
- BNPB. (2013). *Indeks Risiko Bencana Indonesia*. Indonesia: Direktorat Pengurangan Risiko Bencana.
- Buntoro, G. A. (2016). Analisis Sentimen Hatespeech pada Twitter dengan Metode Naive Bayes Classifier dan Support Vector Machine. *Dinamika Informatika*.
- Buntoro, G. A. (2014). Sentiment Analysis Twitter dengan Kombinasi Lexicon Based dan Double Propagation. *ECITEE 2014*, 39-43.
- Buntoro, G. A. (2016). Analisis Sentimen Hatespeech pada Twitter dengan Metode Naive Bayes Classifier dan Support Vector Machine. *Dinamika Informatika*.
- Buntoro, G. A., Adji, T. B., & Purnamasari, A. E. (2014). Sentiment Analysis Twitter dengan Kombinasi Lexicon Based dan Double Propagation. *CITEE 2014*, 39-43.

- Edi Winarko, N. D. (2014). Analisis Sentimen Twitter untuk Teks Bahasa Indonesia dengan Maximum Entropy dan Support Vector machine. *IJCCS*, 91-100.
- Indrayuni, E. (2016). Analisa Sentimen Review Hotel Menggunakan Support Vector Machine Berbasis Particle Swarm Optimization. *EVOLUSI-Jurnak Sains dan Manajemen AMIKBSI Purwokerto*, 4(2).
- Karamizadeh, S., Abdullah, S.M., Halimi, M., Shayan, J., & Javad Rajabi, M. (2014). Advantage and Drawback of Support Vector Machine Functionality. In *Computer, Communications, and Control Technology(I4CT)*, pp. 63-65.
- KemKes. (2017, November 17). *Bencana Alam Yang Terjadi Akibat Faktor Geologi*. Diambil kembali dari Kementerian Kesehatan Republik Indonesia: <http://penanggulangankrisis.kemkes.go.id/bencana-alam-yang-terjadi-akibat-faktor-geologi>
- Munawarah, R., Soesanto, O., & Faisal, M.R. (2016). Penerapan Metode Support Vector Machine pada Diagnosa Hepatitis. *KLIK-KUMPULAN JURNAL ILMU KOMPUTER*, 103-113.
- Priyanti, A, & Ilham, N. (2011). Dampak erupsi gunung merapi terhadap kerugian ekonomi pada usaha peternakan. *WARTAZOA. Buletin Ilmu Peternakan dan Kesehatan Hewan Indonesia*, 21(4), 153-160.
- Putranti, N. D., & Winarko, E (2014). Analisis Sentimen Twitter untuk Teks Bahasa Indonesia dengan Maximum Entropy dan Support Vector machine. *IJCCS*, 91-100.
- Putranti, N. D. (2014). Analisis Sentimen Twitter untuk Teks Berbahasa Indonesia dengan Maximum Entropy dan Support Vector Machine. *IJCCS(Indonesian Journal of Computing and Cybernetics System)*, 8.1:91-100.
- Sabariah, M. K., & Effendy, V. (2015, May). Sentiment analysis on Twitter using the combination of lexicon-based and support vector machine for assessing the performance of a television program. In *Information and Communication Technology (ICoICT), 2015 3rd International Conference on* (pp. 386-390).
-